

DOCUMENT RESUME

ED 172 567

FL 010 401

AUTHOR Henisz-Thompson, Bozena
TITLE A Review of the State of Machine Translation.
PUB DATE [79]
NOTE 26p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Computer Oriented Programs; *Languages;
*Linguistics; *Machine Translation; Programming;
Syntax; Translation; *User Satisfaction
(Information)

ABSTRACT

A review of the state of machine translation considers the historical development of the process, how it works, and results of users' evaluations. Machine translation is the process of translating one language into another through the use of the computer, following the linguistic and computational instructions supplied to it by human researchers in linguistics and programming. The texts to be translated have to be presented to the computer in machine-readable form. Only machine translation that does not involve human editing is considered. The machine translation process that is discussed is unidirectional between a pair of languages: the translation is possible only from Russian to English, for example, and not vice versa, in one system. One of the difficulties that machine translation is faced with is the ability of adjustment to context, but an advantage of machine translation is the speed. An evaluation was conducted by users of the Georgetown system in which translations were unedited translations of Russian texts into English. A second study reviewed evaluations by users of translations made with the FR-II system, involving translations of texts in mathematics, from French into Russian. Organizations that produce machine translations and details of their services and centers conducting research in machine translation are discussed. (SW)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED172567

FL010401

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Bozena Henisz-Thompson

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

A REVIEW OF THE STATE OF MACHINE TRANSLATION

Bozena Henisz-Thompson

INTRODUCTION

Machine Translation is the subject of significant and heated discussions again. Those of us who were involved in its development never doubted that this was bound to happen. It is inconceivable that Machine Translation would die. But it did fade away for a while.

I welcome the opportunity, especially at this time, to recount briefly some of the background of Machine Translation (MT), give some idea of what it involves and how it works, indicate the reactions of its users, and sketch the recent scene.

From the beginnings of MT and throughout its twenty-five-year history, both philosophical attitudes as to its nature and practical solutions for its implementation have varied. But from this retrospective view it is the commonness in the many approaches rather than the differences that is by far the most interesting.

MT, as discussed in this paper, and as understood and implemented by its product-oriented practitioners, is the process of translating one language into another -- for instance, Russian into English or French into Russian -- in which the computer does all of the work involved in the process, following the linguistic and computational instructions supplied to it by human researchers in linguistics and programming. The texts to be translated have to be presented to the computer in machine-readable form; human preparation of the text into such form is the standard practice.

The amount of human participation other than that has always been a controversial issue, usually discussed under the terms of precediting and postediting, especially the latter. These terms refer to human preparation of the texts to be translated in ways that might facilitate translation by machine and correction of the translation

produced by the machine. In this paper, MT refers to translation that does not involve human editing. The field of machine-aided translation in which automated dictionaries and similar helps are used is a rather well-established one with ongoing practical work and research in several centers. But its discussion is beyond the scope of this paper.

The bibliography on MT is already quite extensive, and only a few references to works mentioned in this paper can be given here. The interested reader is referred to the forthcoming book (1) which treats three aspects of MT in considerable detail: historical background, problems and implementation, and users' evaluations.

SOME BACKGROUND

Thoughts that translation could be done by computers clearly preceded an experimental demonstration of the feasibility of such a process. It is the latter, however, that can be taken as a clear-cut starting point. The description of the first experiment in MT is found in the firsthand reports of the coauthors of the experiment (Dostert, 1955; Garvin, 1967). This so-called Joint IBM-Georgetown Experiment of 1954, just under twenty-five years ago, involved the machine translation of a number of sentences from Russian into English. The feasibility of MT was thus established, and the experiment was heralded by the press with fanfare and boundless promises.

But financial support did not follow, and research was continued on a rather limited scale (Macdonald, 1963), primarily at the Georgetown University Machine Translation Project. Matters changed considerably after Soviet researchers announced in 1956 their computer translation from English into Russian and linked it to the Georgetown-IBM experiment. The first experiment was performed in 1955 and described in 1956 (Panov, 1956). Interest and funds suddenly flowed. A great number of meetings and conferences were organized, both national and international; a great many research projects were formed, both in this country and abroad; numerous publications appeared.

The rapid growth of the field and the proliferation of ideas and viewpoints also inevitably led to conflicting views, at times bitter controversies, and more-or-less subtly veiled rivalries. Cooperation

between projects can at best be characterized as low-keyed in spite of voluminous writing and even exchange of research materials.

Despite, however, the perhaps inevitable growing pains of the field, progress was made -- especially by the groups which adopted a product-oriented philosophy and set themselves goals of achieving translations which would be useful in technical fields to serious users. Such attitudes involved what was referred to as "cyclic procedures," which meant gradual improvement of the quality of translations based on experience with actual translations, rather than attempts at global understanding of the languages selected for translation before implementing a working system. Some more notable projects of this kind were the Georgetown MT Research Project, some Soviet groups, and the group of B. Vaugois in France. The work of these projects has resulted in translations which have met the test of users on a substantial scale. Historical accounts of a great number of MT projects and approaches are found in Zarechnak (forthcoming), Macdonald (1963), Josselson (1971), Kulagina and Melcuk (1971), Lehmann and Stachowitz (1971), Vaugois (1976), Katov (1976), Melčuk and Ravič (forthcoming), and others.

By 1961 support for experimental MT started fading, to be supplanted by support for theoretical work in linguistics and mathematical linguistics aimed at establishing complete models of language -- which was considered by some to be a necessary step before MT could be attempted.

An event which had a profound influence on the development of MT was the publication of a small, black-covered book in 1966, usually known as the ALPAC Report (2). The nature of the influence was that of halting altogether or significantly slowing down research and development in MT both in this country and in other countries, although in the latter case the consequences were less drastically felt. ALPAC was the Automatic Language Processing Advisory Committee formed at the request of the National Science Foundation in 1964 in order to evaluate progress and usefulness of MT. It published its report entitled Language and Machines: Computers in Translation and Linguistics in 1966 with this conclusion on page 10: "'Machine translation' presumably means going by algorithm from machine-readable source text to useful target text without recourse to human translation or editing. In this context, there has been no machine translation of general scientific text, and none is

in immediate prospect." The committee considered MT output to be too slow, inaccurate, and not favorably comparable even in terms of cost with human translation, which was considered quite capable of meeting demands for translation. As to the existence of need for MT, the authors concluded that "we find none" (p. 24).

This paper is certainly not the place to go deeper into the above document. It might just be worth mentioning, however, that its negativism is reminiscent of the negative evaluation by Yehoshua Bar-Hillel in 1960 of what he termed Fully Automated High-Quality Translation (FAHQT), whose nonfeasibility he attempted to prove at length (8). It is also worth mentioning that Bar-Hillel modified his attitude very considerably in 1971, as the following quotation attests.

When I started using the term "high quality" a number of years ago, I was using it in a much too absolute sense which cannot be seriously justified. "High quality" has to be relativized with regard to users and with regard to situations. A translation which is of good quality for a certain user in a certain situation might be of lesser quality for the same user in a different situation or for a different user, whether in the same or in a different situation. What is satisfactory for one need not be satisfactory for another. . . . Every program for machine translation should be immediately tested as to its effects on the human user. He is the first and final judge, and it is he who will have to tell whether he is ready to trade quality for speed, and to what degree. [Bar-Hillel, 1971]

The changing attitudes toward MT led to the organization of a conference at the University of Texas in 1971. This conference, attended by foremost experts in the field of MT, computational linguistics and linguistics, had as its purpose the examination of "the controversial issue of a fully automatic high quality translation (FAHQT) in the light of the past and projected advances in linguistic theory and hardware/software capability." This statement is found on page iv in the Technical Evaluation of the report resulting from the conference, entitled Feasibility Study on Fully Automatic High Quality Translation (4). The quotation from Bar-Hillel is contained in his paper in this report and reflects well the opinions of many participants as to the user's role in judgments concerning the quality of translations.

Among the conclusions and recommendations of that report is that "studies should be performed to measure the extent to which comprehensibility of a translation is dependent on the knowledge available to

the actual user" (p. 46). The overall recommendation of the study is that MT should be supported in view of the fact that "quality translation can be achieved in the near future" (p. 48). The major remaining problem is seen to be that of linguistic analysis, since advances in computer hardware (machine capabilities) and software (programs) have removed obstacles in those areas. Wide-ranging research in linguistics is recommended in lexical, syntactic, and semantic areas, including discourse analysis. MT is seen as a field which requires the study of problems beyond the sentence boundary, and thus one which would contribute to the knowledge of language in general.

The years that followed reflect some of these recommendations. Linguistic research developed vigorously, and some attempt was at times made to link it to MT. Mostly, however, it remained on the theoretical level.

Similar history is told for Soviet MT for the years 1961-1974 by Kotov (1976). "This period is characterized by the prevailing development of theoretical studies of language outside of any connection with the specific task of designing practical working MT systems" (p. 6). "The studies and research aimed at designing MT systems as practical and working systems found themselves in quite a different situation. These efforts were not supported and as a result there is not a single working MT system in our country" (p. 7). Interest in undertaking efforts towards a working MT system was shown again, however, in 1973-74 by an official government committee, which concluded that practical MT is feasible, five times cheaper than human, and at least ten times quicker. The committee's recommendation for work towards establishing practical MT without waiting for the solution of all the theoretical problems clearly had an influence. It led to the reevaluation of orientations, which is well evidenced in the conclusions of the 1975 conference on MT in Moscow. These are reviewed under the section on the Recent Scene of this paper.

THE MACHINE TRANSLATION PROCESS

A concise, yet detailed and readable exposition of the problems involved in MT and their solutions is found in Macdonald (forthcoming). A complete documentation of how the Georgetown MT system works is in its

General Report (Macdonald, 1963). The discussion here is extremely general, avoiding details, examples, and keeping technical terms at a minimum. As was mentioned in the Introduction, MT refers to the fully automatic process carried out by the computer of unidirectional translation between a pair of languages. The objection might be raised that it is unreasonable to talk about "the machine translation process" in view of a great variety of approaches and systems that have evolved as a result of them. It is felt, however, that in a description aiming at providing basic understanding of the processes involved what is common in the MT systems should be stressed and brought out, even at the cost of distorting the picture of any one particular system.

MT is distinguished from human translation (HT) by the essential feature of the involvement of the computer. But the computer does not think, and it does only what the human tells it to do. The quality of machine-produced translations is as dependent on the human as is the quality of human translations. However, in MT, human intelligence and knowledge is applied only in the stage of building and improving the system. It is a notorious fact that the quality of human translation varies tremendously. (Actually, as was pointed out by users [Henisz-Dostert, 1973], machine translation is at an advantage in comparison with human translation with respect to what was called "honesty" -- the computer does not paraphrase or invent terms in cases when it does not know how to translate. There are enough similarities between MT and HT to make a relatively simple description of the MT process possible. The part constituting the difference, the computational processes, is only a repetitive application of linguistic knowledge required in both MT system design and human translation.

The MT process sketched here is unidirectional between a pair of languages. This means that given for instance Russian and English, the translation is possible only from Russian to English, and not vice versa, in one system. As is well known to human translators (particularly simultaneous interpreters, where high speed is essential), the set of problems involved in translating from any one given language to another is quite distinct in each case. In MT research the possibility of using some intermediate common language in cases of related languages, for instance the Slavic family, has been investigated to some extent. But

to the knowledge of this author, no such language has been employed. The pair of languages involved in translation are commonly referred to as the Source language, the language of the input text, say, Russian, and the Target language, the language of the output, say, English.

What the human translator starts with (and constantly improves) is the knowledge of the languages with which he or she works, including of course, one's own native language (with respect, for instance, to terminology). This knowledge involves what is commonly referred to as grammar and vocabulary. By a fast and at first poorly understood process, the human mind acquires a vast knowledge of one language (in some cases more than one) at an early stage in life. This knowledge is mostly the acquisition of grammar. To that we add, throughout life, the knowledge of vocabulary, idioms, specialized jargons and styles. Both the human translator and MT system designer make use of this knowledge, but in the translation of specialized texts, such as scientific materials, technical vocabulary and jargon present constant difficulties.

MT processes also involve difficulties which the human does not have. The human's knowledge of language is intimately tied with understanding of the environment, and thus provides the ability to adjust to circumstances, general as well as those of linguistic context. One of the difficulties that MT is faced with is this ability of adjustment to context. Specialized, and particularly scientific, texts offer somewhat restricted style and context. It is thus not surprising that the successful user-tested machine translations have been translations of scientific texts.

MT has advantages, however, over the human translator -- it can be considerably faster, and its user is inclined to be much more tolerant of bad grammar and style. Speed is a very important consideration, of course. But also, the users of MT accept from the machine what they would not accept as easily from a human. They are willing to apply their own knowledge of context to compensate for the machine's lack of adjustment.

Just as in the case of the human translator, the MT process has to involve the knowledge of the grammar and the vocabulary of both languages translated; the building of that knowledge is primarily the task of the linguistic researcher. Secondly, this knowledge has to be

presented to the computer in the form of suitable programs. This task rests primarily with the researcher trained in programming. But there seldom can be a complete separation of these roles, and thus a new breed of scientists, computational linguists, came into being. However, a large degree of specialization of the two roles, the linguist and the programmer, still persists in the majority of centers involved in computational linguistics, to which machine translation gave rise and of which it is now a part.

The knowledge of the grammar and the vocabulary for the MT process cannot be stated in the form of conventional grammar books and dictionaries, since in this form it could not be translated into instructions for computational operations, and would contain inadequate information anyway. The majority of words in one's own native or other well-known language are seldom looked up in a dictionary for all the information about them. In the MT process, the dictionary is constantly consulted and the information contained in it tends to be extensive. In the case of a morphologically rich language, for instance Russian, the dictionary includes information which is the result of detailed morphological analysis. Thus, the information for a noun may include its animateness, gender, number, case endings with which it can be used (and which may require the insertion of some preposition in the translation), and of course its equivalent in the target language. The lexical equivalent in the target language must appear, clearly, in the appropriate form. Thus, in the case of English nouns, they would be singular or plural, and possibly possessive. If the target language is a highly inflected language such as Russian, extensive morphological processes, sometimes called morphological synthesis, are involved in building the output lexical item. The choice of an adequate equivalent is no easy matter since the translation of words often varies depending on context. Special complex processes must be developed and employed which select equivalents upon sensing the context. Provisions have to be made for dealing with words that are not in the dictionary. Morphological analysis is of help here, so that a given word can be recognized as being, for instance, a noun in a certain case that requires the insertion of a preposition in the target language. Such a word appears in the translation in its source language form since no equivalent is available, but

its function may be signaled by the results of morphological analysis, for instance the insertion of a preposition. An important part of lexical analysis is the treatment of idioms. An idiom is considered to be any segment the meaning of which is not the sum of the meanings of its parts. Thus a segment such as "hot dog" is usually an idiom. The translation process is greatly facilitated if such segments can be recognized and dealt with first rather than being subjected to a dictionary search. Thus, the analysis of idioms may be the first step in scanning the source text by means of consulting a specially constructed dictionary of idioms.

When dictionary lookup is completed, a good deal of information is already available about the lexical items. But a sentence is clearly not just a bunch of words. Syntactic analysis is therefore necessary to ascertain their relations and functions. Unfortunately, linguistic knowledge of these relations and functions is still inadequate. This lack of understanding is felt acutely in the MT process for which the rules have to be formulated in extremely precise ways. A grammar full of exceptions, restrictions, and special cases is not very useful for MT. Various approaches to syntactic analysis have evolved from rather simple phrase-recognition techniques to transformational analysis involving the identification of deep structures. For a language such as Russian and many others, identification of structures of agreement and government is extremely useful. For this reason, many researchers use various forms of dependency grammar (Hays, 1964) which establishes such syntactic relations. The major problem in the analysis of a sentence involves the recognition of the subject group and the predicate group, and in languages in which the word order is relatively free -- of which Russian is again a good example -- this is no trivial matter. If the subject is not in the nominative case, the problems can clearly be severe. A great number of "smaller" problems, for example negative structures, also need detailed analysis (for instance, Russian employs multiple negatives and English does not).

That syntactic organization is different in different languages is known even to most laymen. Researchers in contrastive linguistics have conducted many studies comparing syntactic structures in a number of languages. In the MT process, such problems of comparison and transfer

play a central role. Clearly, if a sentence is not correctly analyzed in the first place, it will not be correctly transferred. But even if analysis is successful and correct, the problem of establishing appropriate syntactic organization in the target language is severe. This process of establishing the correct order of final output sentences is sometimes called syntactic synthesis. Its relation to syntactic analysis sometimes resembles that of a mirror image, since very extensive reorganization of the sentence often needs to take place in the process of transfer.

The problem of dealing with units larger than sentences, for instance paragraphs, has received some attention in MT research, but has not been part of operational MT. This problem of context beyond a single sentence in general is quite an important one -- for instance in solving the ambiguities which pronouns may present with respect to their referents.

The various steps in the MT process can be summarized in terms of the description which the Soviet scientist Olga Kulagina provided of her FR-II system at the Seventh International Conference on Computational Linguistics (Coling, 1978).

The FR-II system is designed for translating French into Russian. Its field of application is texts in mathematics. This restriction to one field makes it possible to maintain the dictionary of word stems (morphological analysis of endings is done separately for complete lexical items) at the rather low level of about 1,800. The system also includes a dictionary of about 300 idioms.

The system has the following components:

- (1) Dictionary of French bases
- (2) Dictionary of idioms
- (3) Morphological analysis and treatment of unknown words
- (4) Resolution of homography¹ with the help of context
- (5) Syntactic analysis (based on dependency grammar)
- (6) Transformations
- (7) Syntactic synthesis
- (8) Morphological synthesis

The design of Kulagina's system is very reminiscent of the Georgetown MT system. For the sake of comparison, the size of the Georgetown dictionary was much larger, about 40,000 items, and the dictionary covered eight fields in 1963. The Russian equivalents in the Georgetown dictionary were derived from recent translations of actual texts in those disciplines, thus attempting to keep the MT dictionary

abreast of current technical vocabulary. When the Georgetown system was used at EURATOM (the European Common Market Commission Research Center) and at the United States Atomic Energy Commission's Oak Ridge Laboratory (Henisz-Dostert, 1973), additions of a few thousand words were made at each center. In the United States Air Force use of the Georgetown system the size of the dictionary was increased to about 150,000 items.

STUDIES OF USERS' EVALUATIONS

This section covers two studies of users' evaluations of machine translations. What these studies have in common is that the evaluations are by users of systems that are actually working. One study was conducted by the present author (Henisz-Dostert, 1973), the other by O. Kulagina, mentioned in the previous section, in 1977.² The 1973 study was based on the evaluations by users of the Georgetown system in the years 1963-1973 which operated at two major centers: EURATOM in Italy, and at the United States Atomic Energy Commission's Oak Ridge National Laboratory. The translations were unedited translations of Russian texts, mostly articles, into English made at the request of scientists in a great number of fields -- over thirty, as reported by users, with a dictionary that was by no means suited to many of them. The 1977 study reviewed evaluations by users of translations made with the FR-II system. These were translations of texts in mathematics, from French into Russian. The 1973 study was based on evaluations by fifty-eight users in a variety of fields, the 1977 one on those by one hundred users who were either mathematicians or linguists. The number of pages of translations on which the users' evaluations are based in the 1973 study is quoted by the users themselves to be just under 14,000 (or about 5.5 million words, assuming a printed page of 400 words). Such data is not available to me for the 1977 study, but the numbers can be expected to be smaller.

The controversy between Bar-Hillel and M. Zarechnak reported on in the 1971 Feasibility Study (4) on the quality of MT showed well that a common criterion for such judgments is almost impossible. While Bar-Hillel claimed that MT output was only 35-40 percent complete, Zarechnak put the figure at 80 percent. But, as was pointed out, Zarechnak's figure, based on reactions of scientists, referred to

informativeness of translations, while Bar-Hillel's to their readability. Even these terms are difficult to achieve a consensus on, and probably carry somewhat different meanings both for the researchers and the users. If quite distinct terms are used in evaluative studies, comparisons are made even more difficult. In the case of the two studies compared here, this difficulty does not, however, seem to be major. Yet it is considered necessary to quote the summary results independently before attempting comparisons.

The 1973 study contained two questions which were aimed at measuring the informativeness of MT output. These were:

- Can you decide on the basis of a machine translated text whether the work described is important to you?
- Do you get enough information about the work described?

The answers to these were 100 percent for the former and 85.7 percent for the latter, which yields 92.9 percent as an index of informativeness of MT output.

As for the quality, the question was:

- How do you rate the quality of MT in general?

The evaluations are summed up as:

Good	15.5%
Acceptable	75.9%
Poor	8.6%

The two top categories thus yield 91.4 percent, which can be taken as a measure of adequacy of MT texts.

Readability may be measured in terms of the results of the question which inquired about correctness of translation of sentences. The table which follows gives these results.

<u>SENTENCES</u>	<u>MEAN</u>
Correct	17.1
Understandable	49.2
Understandable with difficulty	20.9
Badly distorted	9.7
Unintelligible	5.39

What these figures show is that 66.3 percent of the sentences are in the combined group of "correct" or "understandable," and 87.2 in the first three categories. The figures can be equated with grammatical correctness, since English morphology involved is negligible.

Kulagina presented her results in the following table:

<u>BIEN [Good]</u>	<u>MOYEN [Medium]</u>	<u>MAL [Bad]</u>
Compréhensibilité [understandability]		
63.6%	24.9%	11.5%
Correction grammaticale [grammatical correctness]		
53.5%	32.7%	13.8%
Adéquation [adequacy]		
66.7%	22.5%	20.8%

It is obviously not too easy to compare these figures with the ones obtained in the 1973 study, the major three of which are:

Adequacy	91.4%
Informativeness	92.9%
Readability (sentences)	87.2%

The overall picture, however, emerges rather clearly. If the first two categories in Kulagina's figures are combined, which for purposes of comparison seems reasonable, her figures would be:

Understandability	88.5%
Grammatical correctness	86.2%
Adequacy	89.2%

The highest figures appear in both studies for adequacy, informativeness, understandability of MT texts. It is not surprising that grammatical correctness is lower, especially if only top evaluations are considered. But what the results indicate beyond any doubt is that studies of actual users' evaluations put the figure of usability of machine translations at close to the 90 percent level.

THE RECENT SCENE

The most satisfying way to start this section would be by listing organizations which produce MT, giving details as to the languages handled, fields, cost, speed, and, possibly, user ratings. Next, one would like to discuss a number of centers conducting research in MT aimed both at higher quality of translations and furthering knowledge about language in general. Hopefully, we would be able to talk not only about individual achievements, but also cooperation and coordination of efforts, possibly internationally. At present, what can be said in both regards is quite unsatisfactory. But

the interest in MT (and in computer-aided translation) has been on a definite increase, both in the practical applications and theoretical research related to MT, in large part as a result of the growing need for vast volumes of translation, the need which can no longer be met by conventional human translation. Both the interest and need has been manifest in recent discussions of MT at government-sponsored seminars and conferences of researchers in this or related fields. In the U.S., a Seminar on Machine Translation was organized at the request of the U.S. Government Foreign Broadcast Information Service (FBIS) at Rosslyn, Virginia in March 1976 (18) which gathered about 100 people who were either contributors or listeners. A number of approaches and efforts, both practical and theoretical, were discussed, and some are reviewed here. This was followed by a seminar in May 1978, but it is as yet unclear to what extent MT (as distinguished from machine-aided translation) was the subject of consideration, although persons currently actively engaged in MT were consulted. In the Soviet Union, a conference was held in Moscow in late 1975 which formulated directives for the development of practical MT, and evidence is given here that they have been followed. At the Seventh International Conference on Computational Linguistics (COLING 78) at Bergen, Norway in August 1978, a section of papers and sessions were devoted to Mechanical Translation in which the researchers from several countries discussed their recent results and approaches in practical systems, theoretical work and computer-aided systems. Also, a full-house panel session was organized entitled Computer Aids to Translation, which included considerations of MT. This conference, which I attended, constitutes the basis for several discussions in this article.

The recent scene can be roughly characterized as consisting of

three foci: (1) systems actually producing usable translations; (2) research on systems (especially of the Artificial Intelligence type) aimed at MT; (3) work on computer-aided translation. The discussion here centers on (1), attention is given to (2), and only a mention made of (3).

With respect to (1), in this country, the Georgetown system is routinely used to produce translations at the request of scientists at the Oak Ridge National Laboratory. It has been in use there since 1965 and had in the meantime undergone reprogramming for the IBM 360/70 computers. In 1976 Martha Gerrard of ORNL reported that 75 percent of all translations made there were computer-produced, and that the remaining 25 percent of texts were usually unsuitable for machine processing (18). The size of the dictionary was reported increased to about 50,000 items. While the cost was still comparable to the cost of human translation, the speed of MT was said to be its greatest advantage, since the translation of a ten-page article can be delivered to the requestor in a day. The Georgetown System is currently running at Georgetown University itself, but so far is used mainly for experimental purposes. Linguistic improvements are being formulated and incorporated into the system. There have been indications of interest in its use by some firms and agencies following a recent demonstration (personal communication, 1978).

At the FBIS conference some commercially oriented systems which are basically derived from the Georgetown System were presented, but as pointed out by S. Petrick in his summary of the conference, they are only marginally different from their predecessors (18).

To continue the discussion of practical systems, the conference in Moscow in 1975 noted that "linguists are beginning again to do more and more practical MT research. The general direction of the research

aimed at the creation of practical MT systems should be based on the actual available experience in formalization of the description of language data and designing of algorithms and programs, and should not be based on untested hypotheses and theories. The investigations and designs should adhere to the following principles:

(1) Orientation of translation systems for a pair of languages, for individual sublanguages, and if necessary, for individual sets of information documents.

(2) Study and description of specific sublanguages with the degree of detail which is actually necessary and could be used in algorithms.

(3) General use of statistical data concerning the distribution of linguistic units in texts taking into consideration their environment and the properties of the linguistic structures in building dictionaries and algorithms" (Kotov, 1976).

These recommendations are well reflected in Kulagina's FR-II system, discussed earlier in this paper. The system deals with a pair of languages (French to Russian); it handles the sublanguage of mathematics; and pays careful attention to units of text specific to the particular documents (it includes a dictionary of idioms whose size is one-sixth of the total dictionary). Also, as was described earlier, the system was put to the test of users, which certainly shows its practical orientation.

In bilingual and language-conscious Canada, the problem of translation has been acute for a long time. In the words of Brian Harris at COLING 78, the cost of human translation there has shot up to 40¢ per word (it was still quoted as being 3¢ per word in the U.S.). The extremely large volumes of translation, high prices of human translation, and government sponsorship created an atmosphere in which practical MT

developed on a large scale and has been routinely used for supplying a large public with information. As was pointed out at COLING 78, an earlier attempt to use MT for translating parliament transcripts could hardly meet with success due to the complexity of the texts. However, a weather-reporting system called METEO was highly successful. This system was described by John Chandious in 1976 (18) and referred to at COLING 78. The system was developed at the University of Montreal and has been in daily use since 1976 for country-wide public weather forecasts. The translation is fully automatic from English to French, and the success rate is 80 percent of sentences (the remaining 20 percent of sentences rejected by the system are translated by humans). The system handles 30,000 words per day, translating with the speed of 1,000 words per minute at a cost of 3.5¢ per word. It includes (1) the idiom dictionary of about 300 entries, (2) the main dictionary (including morphological variations), (3) the parser, whose syntactic analysis is geared toward the semantic microworld of meteorology, and (4) the generator, which synthesizes the French output (including morphology). At COLING 78, P. Isabelle (20) described another system, AVIATION, which the University of Montreal is developing for the government. This system is designed for the translation of technical manuals in aeronautics, with a dictionary of about 50,000 entries. Like METEO, the system is part of the TAUM project, whose long-range aim is fully automatic translation of general texts from English into French. The design bears resemblance to METEO, and as can be easily seen, to Kulagina's system, which, in turn, is reminiscent of the Georgetown System. The translation process is segmented into: (1) morphology of English, (2) dictionary, (3) English analysis, (4) transfer from English to French, (5) synthesis of French, (6) morphology of French.

In France, B. Vauguois' work at Grenoble has always been oriented toward practical purposes. Some of the work of his group was presented at COLING 78, and he himself delivered a paper aimed at a generalized picture of their system with relation to the overall MT scene. The most salient was Vauguois' response to a question from the audience enquiring whether their system design was a sufficient one. The response was, "No, it is not enough, but enough of a solution for important customers," and enough to "have some assurance that ambitious plans can be met." Unfortunately, Vauguois did not elaborate on who the "important customers" are or on the volume of translation. In view of the existence and expansion of the European Common Market and a variety of forms of cooperation within the European Community, one can easily visualize the needs. A paper by N. Nédsbojkiné of the Grenoble group presented at COLING 78 entitled "Etat Actuel du Systeme de Traduction Automatique Multilingue" (Present Status of a Multilingual Automatic Translation System) dealt with the components of their Russian to French automatic translation system. It is therefore interesting to compare it to the organization of Kulagina's French-to-Russian system. The similarities are striking, since the components listed by Nédsbojkiné are: morphological analysis, syntactic analysis, lexical transfer, syntactic transfer, syntactic generation, morphological generation. Also described is an interactive monitor which allows the nonspecialist user easy access to the system.

Turning now to the more theoretical efforts, one finds W. Chafe's (18) characterization of the machine translation model which stresses the importance of forming a conceptual structure of the source language through some process of comprehension of the surface structure; the need for converting that conceptual structure to a conceptual structure adequate for the target language; and the necessity of converting those through some process of verbalization into the surface structures of the target language. A model such as this requires extensive contextual and semantic analysis for the purposes of comprehension of the input text, the understanding of conceptual structures, and the formulation of verbalization processes. Chafe himself turned his research toward the latter upon concluding that real MT takes such deep knowledge that it is utopian. The requirements presented by such a model are indeed enormous. Inputs which attempt solutions come from projects which are usually characterized

as artificial intelligence (AI). At the core of such research is the understanding of texts. The position is that since understanding texts requires not only familiarity with syntactic structures and lexical items, but also the knowledge of general situational context, this knowledge must be built into systems that aim at processing texts automatically. At the 1976 FIBS seminar, Roger Schank of Yale University presented the workings of a program for understanding stories (18). The program maps English into conceptual representation, knowledge of context is utilized to supply missing or implicit inferences about a situation, and references about things contained in a text can be consulted to aid understanding. Once a text is understood, that information can be stated in a language, whether it be a restatement in English or translation by paraphrase into other languages. Chinese, Russian, Dutch and Spanish are mentioned. At the same seminar, Yorick Wilks of the University of Edinburgh stressed the importance of world knowledge in MT and presented "a sketch of a small research English-French MT system that takes in paragraphs on-line and translates them via an interlingua of deep meaning structures and inference rules" (18).

Petrick, in his summary of the 1976 FIBS seminar, concludes that AI systems for natural language understanding "pay a price . . . for their insistence on more adequate conceptual structures, because those structures are not easily obtained for unrestricted text input. . . . Due to the fact that most source language sentences are not processed by these systems, they are unsuitable for unrestricted text and have been applied only to question answering systems and to restricted toy-world domains." Vauguois voiced a similar position in 1976: "All experiments concluded within this artificial intelligence approach are restricted to 'microworlds;' it is too early to consider seriously a generalized use of this method, given the amount of information about the world needed for a real translation of large amounts of text" (6).

At COLING 78, J. Carbonell of the Yale University AI Project described in some detail the approach of that group in his "Towards Knowledge-Based Machine Translation" (20). The stated aim is not to "simulate a professional translator" but rather a "person with a working knowledge of two languages who is asked to read a text and then to reproduce its content in some desired level of detail." This aim

characterizes well the end product: the translations are paraphrases rather than what is conventionally understood by translation. This may be bothersome or downright unacceptable in many applications, but the more troubling part of the stated aim is that concerning "some desired level of detail," since this is not only vague as it stands, but is likely always to be vague or even unstatable. I do not think there are or can be any general standards by which we can establish "the desired level of detail," since that always depends on the user of the translation. The system as outlined by Carbonell works as follows: The source text is analyzed and mapped into a language-free conceptual representation. This requires in-depth understanding of a passage in its linguistic context and the knowledge of the cultural situation. For translation, this knowledge context must be shared by the speakers of the source and target language. Various inference mechanisms apply conceptual world knowledge to augment the representation with information that was implicit in the input text. Knowledge of intentions, for instance, must also be included. Once this knowledge representation is achieved, a language-free summary or paraphrase is formed. Finally, a natural-language generator maps the most relevant sections of the language-free representation into the target language. This system utilizes the program for understanding stories called SAM, developed by Roger Schank. Carbonell states that summarizers and generators of varying degrees of sophistication have been built for Spanish, Mandarin Chinese, Russian, Arabic and Dutch, with the Spanish being the most advanced, and stresses the point that knowledge-based MT is now possible to a limited, but nontrivial, extent.

The AI approach is certainly used and advanced by researchers with much linguistic knowledge and knowledge of computer operation. Apart from the objection that paraphrasing in translation may often be quite unacceptable, two others arise. One is that voiced by Petrick and Vaugois about the limitation of the approach to microworlds. Another is the finding from studies of users' reactions to MT that the knowledge necessary to understand MT texts is supplied by the users themselves (Henisz-Dostert, 1973). But the most serious objection is that the approach of producing paraphrases may be badly misused by persons attempting to develop commercial translations who do not have the linguistic

sophistication that the proponents of the AI approach have, nor the necessary scientific integrity.

Another theoretically oriented system called SALAT -- System for Automatic Language Analysis and Translation -- was presented at COLING 78 by Christa Hauenschild of the University of Heidelberg. The project is called Machine Translation and is at the Institute of Applied Linguistics, but it aims at theoretical rather than practical output. It is being tested on simple sample sentences in German, English, French, and Russian, but is essentially language-free, employing an interlingua which serves as a representation of meanings as well as a knowledge of situation and the world. Special deduction rules manipulate these deep structure interlingua representations. The components of the system are (1) analysis, which may map an input sentence into several deep structure representations; (2) transformational analysis operating on predicates; (3) disambiguation (4) transfer; (5) transformational synthesis. The general orientation of this project is similar to the AI type MT research in the longer term perspective.

A FEW CONCLUSIONS

It would be premature to attempt to judge what directions the development of MT will take. Hopefully, it will be towards practical and empirical approaches joined with theoretical longer-range efforts. In the past, this did not fare very well. But the time is now to push forward in those challenging encounters of man's inadequate knowledge of his most expressive tool -- language -- with his perhaps most daring invention -- the computer. Machine translation forces us to meet that challenge in a number of areas. It forces us to a greater awareness of our mental capabilities in the environment in which humanly-created artifacts assume an increasing role.

FOOTNOTES

1. Homographs are words which are spelled the same but which may have different meanings or may require different equivalents in translation. Such words are responsible for the famous examples of bad machine translations, some of which are purely jocular inventions. Here belongs the notorious translation of the idiom "out of sight, out of mind" as "invisible idiot," and the saying "the spirit is willing, but the flesh is weak" as "the wine is good, but the meat is spoiled."

But there have also occurred some funny real mistranslations due to the inability of the system to sense the context. In this category are such translations as "waterfall reactions" instead of "cascade reactions" in physics contexts; "squirrels circulating in the blood" instead of "albumins circulating in the blood" in a biological context; "volatile mouse" instead of "bat," "Leonardo yes Vinci" instead of "Leonardo da Vinci."

Such mistranslations are in some cases due to one more general meaning being included in the dictionary, such as "waterfall"; in others they are due to intricacies of case endings, such as "squirrels," because in certain cases and genders, two different words in Russian are spelled identically; yet in others, such as "volatile mouse," they are due to idioms and context. Even though in Russian "bat" is called "flying mouse," in translation this has to be treated as an idiom. In the unlikely case that the text would really talk about flying mice, mistranslations would again occur. In the case of Leonardo, "da" got translated as "yes."

It is worth noting that the "waterfall" and "squirrels" examples were pointed out to me by users always with a chuckle, and in the case of "waterfall" even with the comment: "Don't change that; we like it."

2. These studies are reported in detail in forthcoming books by these authors.

REFERENCES

1. Henisz-Dostert, Bozena; Macdonald, R. Ross; and Zarechnak, Michael. Machine Translation. The Hague: Mouton, in press.
2. Language and Machines; Computers in Translation and Linguistics. ALPAC (Automatic Language Processing Advisory Committee) Report, National Academy of Sciences, National Research Council Publication 1416, Washington, 1966.
3. Garvin, Paul. "The Georgetown-IBM Experiment of 1954: An Evaluation in Retrospect," Papers in Linguistics in Honor of Léon Dostert, edited by W. M. Austin. The Hague: Mouton, 1967, pp. 46-56.
4. Lehmann, Winfred and Stachowitz, Ralph. Feasibility Study of Fully Automatic High Quality Translation, Report No. RADC-TR-71-295, Austin, Texas, 1971.
5. Dostert, Léon E. "The Georgetown-IBM Experiment," Machine Translation of Languages, edited by W. N. Locke and A. D. Booth. New York: John Wiley, 1955, pp. 124-35.
6. Vauguois, B. "Automatic Translation -- A Survey of Different Approaches," Statistical Methods in Linguistics, 1976.
7. Josselson, Harry H. "Automatic Translation of Languages since 1960: A Linguist's View," Advances in Computers, 1971, pp. 1-53.
8. Bar-Hillel, Y. "The Present Status of Automatic Translation of Languages," Advances in Computers, 1960, pp. 91-163.
9. _____. "Some Reflections on the Present Outlook of Fully Automatic High Quality Translation," Feasibility Study of Fully Automatic High Quality Translation (see reference 4), pp. 72-76.

10. Kotov, R. G. "Lingvistika i sovremennoe sostojanie mašinnogo perevada v strane (Linguistics and the Contemporary State of Machine Translation in the U.S.S.R.), Voprosy Yazykoznanija, 1976, pp. 37-49.
11. Macdonald, R. Ross. General Report, 1952-1963, Georgetown University Machine Translation Project, Paper No. 30, Washington, D.C., 1963.
12. Henisz-Dostert, Bozena. Users' Evaluation of Machine Translation: Georgetown MT System, Report No. RAD-TR-73-239, Austin, Texas, 1973.
13. Hays, David G. "Dependency Theory: Formalism and Some Observations," Language 40, 1964, pp. 511-25.
14. Kulagina, O. S. and Melcuk, I. A. Avtomaticeskij Perevod (Automatic Translation). Moskva: Progress, 1971.
15. Panov, D. Ju; Ljapunov, A. A.; and Muxin, I. S. Avtomatizacija perevoda s odnogo jazyka na drugoj (Automatization of Translation from One Language to Another). Moskva, 1956.
16. Carbonell, J. G.; Cullingford, R. E.; and German, A. V. "Towards Knowledge-Based Machine Translation," Abstracts, COLING 78, Seventh International Conference on Computational Linguistics, University of Bergen, Norway, 1978 (reference 20).
17. Isabelle, P. et al. "TAUM-AVIATION: un Systeme pour la traduction automatique de manuels techniques" (TAUM-AVIATION: System for Automatic Translation of Technical Manuals), (see reference 16).
18. FIBS Seminar on Machine Translation. American Journal of Computational Linguistics, microfiche 46.
19. Melčuk, I. and Ravič, R. Automatic Translation, 1964-1970, A Bibliography. Department of Linguistics and Philology, University of Montreal, in preparation.

20. COLING 78 Proceedings. Available from K. Heggstad, Nordisk Institutt, PDS, H. Harfagresgt. 29, N-5014, Bergen University, Norway.